STAT 621 Chapter 3

Zepu Zhang
October 5, 2011

1 The Binomial test

Highlights:

1. Recognize the generality and broad applicability of Binomial-based tests.

2. The test statistic (T) and its null distribution.

3. Normal approximation to binomial.

   - "large sample size"
   - "correction for continuity"

4. Alternative hypotheses: two-tailed vs one-tailed.

5. Implications of the discrete nature of T:

   - determination of the rejection region
   - the resultant level of significance (α)
   - to include or not to include "or equal to"

6. How to determine the p-value.

Example Ex. 1, page 127. (Point to learn: how to state (choose) H_0 and H_1.)

Example Ex. 2, page 128.

2 Confidence interval for a probability or population proportion (p)

Highlights:

1. Method A (exact) vs Method B (approx).

2. Derivation of the formulas for Method B.

3. How was the table for Method A made?

Example Ex. 3, page 130.

3 The sign test

The sign test is the oldest of all nonparametric tests, and it is a special Binomial test. While a Binomial test tests whether p is a certain value, say p_*, the sign test tests whether p is 0.5.

Specifically, there exist a sample of a bivariate variable: (x_1, y_1), (x_2, y_2),..., (x_n, y_n). We want to test whether X on average tends to be larger than Y.

Highlights:

1. Convert a bivariate problem to a univariate problem—let $Z = X - Y$, then the bivariate sample induces a univariate sample $z_1, \dots, z_n$. The test becomes whether Z tends to be positive or negative.
2. Get the sign ($+$ or $-$) of z_i. Then the test becomes whether there tend to be more $+$'s than $-$'s. Or, in other words, whether $P(+) = 0.5$. Hence this is a Binomial test with $T = \#\{+\}$ and $p_* = 0.5$.
3. Discard ties; in other words, $z = 0$ contains no info about the tendency of the sign of Z.
4. Independent sample: independence between pairs.
5. "Internally consistent": if $P(+) > P(-)$ for one pair, then this is the case for all pairs. This is one of the requirement of a Binomial experiment: constant p in all trials.
6. How to state the hypothesis: $P(+) = P(-)$ vs alternatives.
7. An alternative statement of hypothesis: $E(X) = E(Y)$ vs alternatives.

   This involves some assumption about the distribution of $Z = X - Y$. Specifically, it assumes the distribution of Z is symmetric about its mean.

4 Variation 1 of the sign test: McNemar test for significance of change

The value of some attribute is one of two categories; let's call them "0" and "1". We take a sample of size n'; some have value "0" and the others have value "1". After a certain incident, the values of this sample is measured again: some will stay the same as before, some “0” switched to “1”, and some “1” switched to “0”. The counts of “0”s and “1”s can be summarized in a contingency table:

		after	
		0	1
before	0	a	b
	1	c	d

The question is: have the fractions of “0”s and “1”s changed?

The test is really whether $b = c$.

Exact test when $b + c \leq 20$: use Binomial (or sign) test.

Large sample test when $b + c > 20$: use normal approx.

Example Ex. 1, page 168.

5 Variation 2 of the sign test: Cox & Stuart test for trend

Example Ex. 2, page 171 (the simple case)

Example Ex. 3, page 171 (filtering out periodicity)

Example Ex. 4, page 172 (detecting correlation)

Example Ex. 5, page 173 (testing for the presence of a pattern of relation)